

Parallel tempering algorithm for integration over Lefschetz thimbles

Masafumi Fukuma* and Naoya Umeda*

Department of Physics, Kyoto University, Kyoto 606-8502, Japan

*E-mail: fukuma@gauge.scphys.kyoto-u.ac.jp; n_umeda@gauge.scphys.kyoto-u.ac.jp

Received March 9, 2017; Revised May 23, 2017; Accepted May 24, 2017; Published July 14, 2017

.....
 The algorithm based on integration over Lefschetz thimbles is a promising method to resolve the sign problem for complex actions. However, this algorithm often meets a difficulty in actual Monte Carlo calculations because the configuration space is not easily explored due to the infinitely high potential barriers between different thimbles. In this paper, we propose to use the flow time of the antiholomorphic gradient flow as an auxiliary variable for the highly multimodal distribution. To illustrate this, we implement the parallel tempering method by taking the flow time as a tempering parameter. In this algorithm, we can take the maximum flow time to be sufficiently large such that the sign problem disappears there, and two separate modes are connected through configurations at small flow times. To exemplify that this algorithm does work, we investigate the $(0 + 1)$ -dimensional massive Thirring model at finite density and show that our algorithm correctly reproduces the analytic results for large flow times such as $T = 2$.

Subject Index A22, B38, D34

1. Introduction

There are many cases where one needs to deal with complex actions. One important example for high energy/nuclear physics is quantum chromodynamics (QCD) at finite density (see Ref. [1] for a review on the recent developments). However, since a complex action does not give a real and positive Boltzmann weight, one cannot directly resort to the traditional Markov chain Monte Carlo methods to estimate correlation functions. The so-called reweighting algorithm (which absorbs the phase of the weight into observables) is highly ineffective when the imaginary part of the action becomes very large (such as in the thermodynamic limit), because one needs to take a sample from a configuration space where the weights of nearby configurations have almost the same amplitudes but very different phases. The difficulty of numerical evaluation in such a situation is termed the *sign problem*.

There have been many proposals to circumvent the sign problem. One of the approaches that are currently under intense study is the use of integration over Lefschetz thimbles [2] (see also Refs. [3–10]), which we will call the Lefschetz thimble algorithm hereafter. There, the original real-valued variable (say, $x = (x^j) \in \mathbb{R}^N$) is complexified according to the *antiholomorphic gradient flow* (sometimes called the *upward flow* in the literature) $\dot{z}^i = [\partial_i S(z)]^*$. In the original Lefschetz thimble algorithm, as will be reviewed in Sect. 2.1, the flow time is taken to infinity to map the original configuration space to a union of Lefschetz thimbles. Since the imaginary part of the action is constant on each thimble, the sign problem disappears if the path integral is made over the Lefschetz thimbles. However, since two different thimbles are separated by an infinitely high potential barrier, one needs to invent some machinery to incorporate contributions from all relevant thimbles.

Recently, Alexandru et al. [11] made a very interesting proposal to consider configurations on a manifold that is obtained from the original configuration manifold by a finite amount of flow time. This algorithm was a great success in various models [11,12], but reducing the amount of flow time may also reduce the effectiveness against the sign problem, and one does not know a priori whether the chosen flow time avoids both the sign problem and the multimodal problem simultaneously.

In this paper, as a versatile tool for Monte Carlo calculations of models with complex actions, we propose a Lefschetz thimble algorithm where the flow time is used as an *auxiliary variable* for the highly multimodal distribution. There can be various methods to realize this idea, and in this paper we implement the parallel tempering method because of its simplicity, by taking the flow time as a tempering parameter. There, we consider a set of manifolds corresponding to various flow times. Two separate modes at large flow times (where the sign problem no longer exists) are then connected by passing through configurations at small flow times (where the original sign problem exists but the multimodality is expected to be mild). Since the sample average is taken only with respect to the largest flow time, we need not worry about the sign problem at small flow times, although we take into account configurations there.

This paper is organized as follows. In Sect. 2 we first review the basics of the Lefschetz thimble algorithm based on Refs. [7,11], and then implement the parallel tempering method in the algorithm by taking the flow time as a tempering parameter. In Sect. 3 we investigate the $(0+1)$ -dimensional massive Thirring model at finite density, and show that our algorithm correctly reproduces the analytic results for large flow times such as $T = 2$. Section 4 is devoted to the conclusion and outlook for future work.

2. Algorithm

2.1. Integration over Lefschetz thimbles (review)

We consider a real N -dimensional dynamical variable, $x = (x^i) \in \mathbb{R}^N$, with action $S(x)$ that may take complex values for real-valued x . Our main concern is to evaluate the expectation values of functions of x :

$$\langle \mathcal{O}(x) \rangle = \frac{\int_{\mathbb{R}^N} dx e^{-S(x)} \mathcal{O}(x)}{\int_{\mathbb{R}^N} dx e^{-S(x)}}. \quad (2.1)$$

We assume that $|e^{-S(x)}|$ decreases rapidly enough in the limit $x \rightarrow \pm \infty$, and that $e^{-S(z)}$ and $\mathcal{O}(z)$ are entire functions when regarded as functions of $z = (z^i) \in \mathbb{C}^N$. The integration region can then be changed to any other region Σ in $\mathbb{C}^N$ as long as it is obtained as a continuous deformation of the original region with the boundary fixed at infinity. We consider as such an integration region the submanifold that is obtained from the following antiholomorphic gradient flow $z(t; x)$ with a flow time t :

$$\frac{dz^i}{dt} = \left[\frac{\partial S(z)}{\partial z^i} \right]^*, \quad z^i|_{t=0} = x^i. \quad (2.2)$$

In fact, the flow defines a map from the original integration region $\Sigma_0 \equiv \mathbb{R}^N$ to a real N -dimensional submanifold Σ_t in $\mathbb{C}^N$:

$$z_t : \Sigma_0 \ni x \mapsto z_t(x) \equiv z(t; x) \in \Sigma_t. \quad (2.3)$$

We thus see that Eq. (2.1) can be rewritten as

$$\langle \mathcal{O}(x) \rangle = \frac{\int_{\Sigma_t} dz e^{-S(z)} \mathcal{O}(z)}{\int_{\Sigma_t} dz e^{-S(z)}}, \quad (2.4)$$

which can be further rewritten as a reweighted integral over $\mathbb{R}^N$ [7] as¹

$$\begin{aligned} \langle \mathcal{O}(x) \rangle &= \frac{\int_{\mathbb{R}^N} dx \det J_t(x) e^{-S(z_t(x))} \mathcal{O}(z_t(x))}{\int_{\mathbb{R}^N} dx \det J_t(x) e^{-S(z_t(x))}} \\ &= \frac{\left\langle e^{i[\arg \det J_t(x) - S_I(z_t(x))]} \mathcal{O}(z_t(x)) \right\rangle_{S_{\text{eff}}}}{\left\langle e^{i[\arg \det J_t(x) - S_I(z_t(x))]} \right\rangle_{S_{\text{eff}}}}. \end{aligned} \quad (2.5)$$

Here, $J_t(x) \equiv (\partial z_t^i(x)/\partial x^j)$ is the Jacobi matrix, $S_R(z_t(x))$ ($S_I(z_t(x))$) is the real (imaginary) part of $S(z_t(x))$, and $\langle * \rangle_{S_{\text{eff}}}$ is the expectation value taken with respect to

$$S_{\text{eff}}(x; t) \equiv S_R(z_t(x)) - \ln |\det J_t(x)|. \quad (2.6)$$

The Jacobi matrix is obtained by solving the combined differential equations with respect to t [7]:²

$$\frac{dz_t^i}{dt} = \left[\frac{\partial S}{\partial z^i}(z_t) \right]^*, \quad z_0^i = x^i, \quad (2.7)$$

$$\frac{d(J_t)_{ij}}{dt} = \left[H_{ik}(z_t) \cdot (J_t)_{kj} \right]^*, \quad (J_0)_{ij} = \delta_{ij}, \quad (2.8)$$

where $H_{ij}(z) \equiv \partial^2 S(x)/\partial z^i \partial z^j$ is the Hesse matrix for the action $S(z)$.

The key point is that the right-hand side of Eq. (2.4) or Eq. (2.5) does not depend on the flow time t , so that one can set t to an arbitrary value that is convenient for actual calculation. Note that under the flow the real part $S_R(z_t(x))$ does not decrease while the imaginary part $S_I(z_t(x))$ is kept constant, because $(d/dt)S(z_t(x)) = |\partial_{z^i} S(z_t(x))|^2 \geq 0$. In the original Lefschetz thimble algorithm, one takes the limit $t \rightarrow \infty$, in which Σ_t approaches a union of connected components (Lefschetz thimbles), and the action has a constant imaginary part on each thimble.³ In a generic situation the phase change coming from $J_t(x)$ is sufficiently mild, so that the Monte Carlo calculation for the expression (2.5) is free from sign problems. However, two different thimbles are also disconnected in the sense of Monte Carlo sampling because S_R increases indefinitely near the boundary of each thimble. This multimodality of distribution makes the Monte Carlo calculation impractical, especially when contributions from more than one thimble are relevant to estimating expectation values.

¹ One may also take as an integration region the tangent space to the critical point of the dominant thimble [7] if the integral can be well approximated by the integration around the critical point.

² The second equation is obtained by differentiating the first equation with respect to $x^j \in \mathbb{R}$;

$$\frac{d(J_t(x))_{ij}}{dt} = \frac{d}{dt} \left(\frac{\partial z_t^i(x)}{\partial x^j} \right) = \frac{\partial}{\partial x^j} \left[\frac{\partial S}{\partial z^i}(z_t(x)) \right]^* = \left[\frac{\partial^2 S}{\partial z^i \partial z^k}(z_t(x)) \frac{\partial z_t^k(x)}{\partial x^j} \right]^* = \left[H_{ik}(z_t(x)) (J_t)_{kj}(x) \right]^*.$$

³ Generically there is a single critical point z_σ on each connected component $\mathcal{J}_\sigma$, and $\mathcal{J}_\sigma$ is obtained as the set of orbits flowing out of z_σ . The complementary submanifold to $\mathcal{J}_\sigma$ in $\mathbb{C}^N$ consists of orbits that flow into z_σ , and will be denoted by $\mathcal{K}_\sigma$. The integrations in Eq. (2.5) are dominated by points near the intersection of $\mathcal{K}_\sigma$ and $\mathbb{R}^N$.

A very interesting proposal made in Ref. [11] is to use a finite amount t , which is chosen to be large enough to avoid the sign problem but also not too large in order to enable exploration in the configuration space. However, one does not know a priori whether the adopted value of t is actually free from the two obstacles (the sign and multimodal problems) simultaneously. We will show that one can solve both simultaneously if we implement the parallel tempering method in their algorithm with the flow time as a tempering parameter.

2.2. Implementation of parallel tempering

The basic idea of the parallel tempering algorithm [13–15] is the following. Suppose that we want to estimate expectation values with action $S(x; \lambda)$, where $x \in \mathbb{R}^N$ is a dynamical variable and λ is the parameter (such as the temperature) that we want to use for a Monte Carlo calculation. The point is that, even when the distribution is multimodal for the original λ (e.g., when λ represents a very low temperature), the multimodality can be made mild if one takes another value $\tilde{\lambda}$ (e.g., $\tilde{\lambda}$ corresponding to a very high temperature). So, if the configuration space is enlarged such that the parameter can change gradually between λ and $\tilde{\lambda}$, two separate configurations for the original λ will be connected by passing through configurations at parameters near $\tilde{\lambda}$. The parallel tempering algorithm enables the move of configurations among different λ by enlarging the configuration space from $\mathbb{R}^N = \{x\}$ to the set of $A + 1$ replicas, $(\mathbb{R}^N)^{A+1} = \{(x_0, x_1, \dots, x_A)\}$. We there assign λ_α to replica α ($\alpha = 0, 1, \dots, A$), such that $\lambda_0 = \tilde{\lambda}$ and $\lambda_A = \lambda$ and that λ_α and $\lambda_{\alpha+1}$ are sufficiently close to each other.⁴ We set up an irreducible, aperiodic Markov chain for the enlarged configuration space such that the probability distribution for $(x_0, x_1, \dots, x_A)$ eventually approaches the equilibrium distribution proportional to

$$\prod_{\alpha} e^{-S(x_{\alpha}; \lambda_{\alpha})}. \quad (2.9)$$

We finally take sample averages only with respect to a sample taken from $\alpha = A$. The simplest algorithm to realize this idea⁵ is to swap two configurations of two adjacent replicas α and $\alpha + 1$, (i.e., to update the configuration $(x_{\alpha} = x, x_{\alpha+1} = x')$ to $(x_{\alpha} = x', x_{\alpha+1} = x)$) with the probability

$$w_{\alpha}(x, x') = \min\left(1, \frac{e^{-S(x'; \lambda_{\alpha}) - S(x; \lambda_{\alpha+1})}}{e^{-S(x; \lambda_{\alpha}) - S(x', \lambda_{\alpha+1})}}\right), \quad (2.10)$$

which obviously satisfies the detailed balance condition

$$w_{\alpha}(x, x') e^{-S(x; \lambda_{\alpha}) - S(x', \lambda_{\alpha+1})} = w_{\alpha}(x', x) e^{-S(x'; \lambda_{\alpha}) - S(x, \lambda_{\alpha+1})}. \quad (2.11)$$

Our proposal is to take the flow time t as such a tempering parameter. The basic algorithm is then as follows.⁶

- **Step 0.** Fix the maximum flow time T , which should be sufficiently large such that the sign problem disappears there, and pick up flow times $\{t_{\alpha}\}$ from the interval $[0, T]$ with $t_0 = 0 <$

⁴ The computational cost required for the parallel tempering method can be roughly estimated to be proportional to $(A + 1)/X$, where X is the minimum of the acceptance rates for all pairs of adjacent replica (see Step 3 below). In this paper, we will set the parameters A and λ_{α} such that the minimum acceptance rate is well above 50% (see Fig. 5).

⁵ Of course, there can be many variations on this algorithm.

⁶ To make discussions simple, we only take the flow time as a tempering parameter. The algorithm can be readily extended such that other parameters are included as extra tempering parameters.

$t_1 < \dots < t_A = T$. The values of A and t_α are determined manually or adaptively to optimize the acceptance rate in Step 3 below.

- **Step 1.** Choose an initial value $x_\alpha \in \mathbb{R}^N$ for each replica α , and numerically solve the differential equations (2.7) and (2.8) to obtain the triplet $(x_\alpha, z_{t_\alpha}, J_{t_\alpha})$.
- **Step 2.** For each α , construct a Metropolis process to update the value of x . Explicitly, we take a value x'_α from x_α using a symmetric proposal distribution, and recalculate the triplet $(x'_\alpha, z'_{t_\alpha}, J'_{t_\alpha})$ using x'_α as the initial value. We then update x_α to x'_α with the probability $\min(1, e^{-\Delta S_{\text{eff},\alpha}})$, where

$$\begin{aligned} \Delta S_{\text{eff},\alpha} &\equiv S_{\text{eff}}(x'_\alpha, t_\alpha) - S_{\text{eff}}(x_\alpha, t_\alpha) \\ &= S_R(z'_{t_\alpha}) - \ln |\det J'_{t_\alpha}| - S_R(z_{t_\alpha}) + \ln |\det J_{t_\alpha}| \end{aligned} \quad (2.12)$$

(recall that $S_{\text{eff}}(x; t) = S_R(z_t(x)) - \ln |\det J_t(x)|$, Eq. (2.6)). We repeat the process sufficiently many times such that local equilibrium is realized for each α .

- **Step 3.** Starting from $\alpha = 0$ through $\alpha = A - 1$, swap the values of x between two adjacent replicas α and $\alpha + 1$ with the probability

$$w_\alpha(x_\alpha, x_{\alpha+1}) \equiv \min \left(1, \frac{e^{-S_{\text{eff}}(x_{\alpha+1}; t_\alpha) - S_{\text{eff}}(x_\alpha; t_{\alpha+1})}}{e^{-S_{\text{eff}}(x_\alpha; t_\alpha) - S_{\text{eff}}(x_{\alpha+1}; t_{\alpha+1})}} \right). \quad (2.13)$$

- **Step 4.** After repeating Steps 2 and 3 sufficiently many times, get a triplet $(x_A, z_{t_A=T}, J_{t_A=T})$ from $\alpha = A$ as an element of a sample.
- **Step 5.** After repeating Steps 2 to 4, we obtain a sequence of triplets $\{(x_A^{(a)}, z_{t_A=T}^{(a)}, J_{t_A=T}^{(a)})\}$, which we use to estimate the expectation value:

$$\langle \mathcal{O}(x) \rangle \approx \frac{\sum_a e^{i[\arg \det J_T^{(a)} - S_I(z_T^{(a)})]} \mathcal{O}(z_T^{(a)})}{\sum_a e^{i[\arg \det J_T^{(a)} - S_I(z_T^{(a)})]}}. \quad (2.14)$$

Note that the action with $t_0 = 0$ is the original action for which the sign problem exists. However, as can be seen from Eq. (2.14), the sample average is taken only with respect to the action with $t_A = T$, so that we need not worry about the original sign problem, although we include configurations near $t_A = 0$. Also, $t_A = T$ can be taken to be sufficiently large such that the sign problem disappears if we complement intermediate flow times sufficiently (with larger A). Thus, this simple algorithm solves the two obstacles simultaneously: the original sign problem at $t_0 = 0$ is resolved at $t_A = T$ while the multimodal problem at $t_A = T$ is resolved by passing through configurations near $t_0 = 0$.⁷

3. Example

In this section we investigate the $(0 + 1)$ -dimensional massive Thirring model at finite density [10,16,17] to exemplify that the algorithm given in the previous section does work.

The $(0 + 1)$ -dimensional massive Thirring model is defined from the standard $(1 + 1)$ -dimensional massive Thirring model by dimensional reduction. With an auxiliary field $\phi(\tau)$, the continuum representation of the grand partition function $Z = \text{tr} e^{-\beta(H - \mu Q)}$ is given by the path integral

$$Z = \int_{\text{PBC}} [d\phi(\tau)] \int_{\text{ABC}} [d\bar{\psi}(\tau) d\psi(\tau)] e^{-S[\phi, \bar{\psi}, \psi]}, \quad (3.1)$$

⁷ We have implicitly assumed that the action at $t_0 = 0$ does not cause a multimodality in the configuration space. If this is not the case, we further introduce other parameters (such as the coefficient of the action) as extra tempering parameters or prepare flow times $\{t_\alpha\}$ such that $t_0 < 0$.

where the Euclidean action takes the form

$$S[\phi, \bar{\psi}, \psi] = \int_0^\beta d\tau \left[-\bar{\psi} (\gamma^0 (\partial_0 + i\phi + \mu) + m) \psi + \frac{1}{2g^2} \phi^2 \right], \quad \left[\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right], \quad (3.2)$$

and the ϕ integral (the ψ , $\bar{\psi}$ integral) obeys the periodic (antiperiodic) boundary condition. In realizing the model on the lattice, we discretize the Euclidean time as $\tau = na$ ($n = 1, \dots, N$) with $\beta = Na$ (N : even), and follow the prescription of Ref. [17], where $\phi(\tau) = \phi_n$ is treated as a $U(1)$ gauge potential and is combined with the chemical potential μ to become a link variable of a complexified gauge group, $e^{(i\phi(\tau) + \mu)a} = e^{i\phi_n a} e^{\mu a} \equiv U_n e^{\mu a}$, as proposed in Ref. [18]. Then, by using a staggered fermion formulation, the grand partition function is given by

$$Z = \int (dU) e^{-S(U, \bar{\chi}, \chi)}, \quad (3.3)$$

where $(dU) = \prod_n (dU_n) \equiv \prod_n [d(\phi_n a)/2\pi]$ (with $\phi_n a \in (-\pi, \pi]$) and

$$S(U, \bar{\chi}, \chi) = - \sum_{n, \ell=1}^N \bar{\chi}_n D_{n\ell}(U) \chi_\ell + \frac{1}{2g^2 a} \sum_{n=1}^N \left[1 - \frac{1}{2} (U_n + U_n^{-1}) \right] \quad (3.4)$$

with $D_{n\ell}(U) \equiv (1/2) (U_n e^{\mu a} \delta_{n+1, \ell} - U_{n-1}^{-1} e^{-\mu a} \delta_{n-1, \ell} - U_n e^{\mu a} \delta_{n, N} \delta_{\ell, 1} + U_N^{-1} e^{-\mu a} \delta_{n, 1} \delta_{\ell, N}) + m a \delta_{n\ell}$. We henceforth set $a = 1$ and treat m , μ , and g^2 as dimensionless parameters.

After carrying out the fermion integration, we obtain

$$Z = \int (dU) \det D(U) e^{-(1/2g^2) \sum_n [1 - (1/2)(U_n + U_n^{-1})]}, \quad (3.5)$$

which can be calculated analytically [17] as

$$Z = \frac{e^{-N\alpha}}{2^{N-1}} [\cosh(N\mu) I_1^N(\alpha) + \rho_+ I_0^N(\alpha)], \quad (3.6)$$

where $\alpha \equiv 1/(2g^2)$, $2\rho_\pm \equiv (\sqrt{m^2 + 1} + m)^N \pm (\sqrt{m^2 + 1} - m)^N$, and $I_n(\alpha)$ is the modified Bessel function of the first kind of order n . Using this expression, we obtain the analytic form of the chiral condensate as

$$\langle \bar{\chi} \chi \rangle = \frac{1}{N} \frac{\partial \ln Z}{\partial m} = \frac{\rho_- I_0^N(\alpha)}{\sqrt{m^2 + 1} [\cosh(N\mu) I_1^N(\alpha) + \rho_+ I_0^N(\alpha)]}. \quad (3.7)$$

We now numerically evaluate the chiral condensate by using the complex action

$$S(U) \equiv \frac{1}{2g^2} \sum_n \left[1 - \frac{1}{2} (U_n + U_n^{-1}) \right] - \ln \det D(U) \quad (3.8)$$

as

$$\langle \bar{\chi} \chi \rangle = \left\langle \frac{1}{N} \text{tr} \left(D^{-1}(U) \frac{\partial D(U)}{\partial m} \right) \right\rangle = \left\langle \frac{1}{N} \text{tr} D^{-1}(U) \right\rangle, \quad (3.9)$$

where

$$\langle \mathcal{O}(U) \rangle \equiv \frac{\int (dU) e^{-S(U)} \mathcal{O}(U)}{\int (dU) e^{-S(U)}}. \quad (3.10)$$

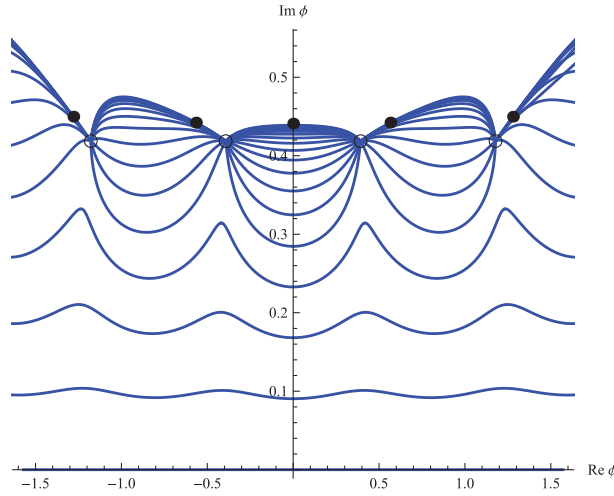


Fig. 1. $\phi(t) = (1/N) \sum_n \phi_n(t)$ at flow times $t = 0, 0.1, 0.2, \dots, 1.9, 2.0$ from bottom to top. The parameters are set to $N = 8, m = 1, \mu = 1.3, g^2 = 1/2$. The full circles are the critical points of S , and the empty circles are the log singularities of S .

It is easy to check that $[\det D(U; \mu)]^* = \det D(U; -\mu)$, so that the second term of Eq. (3.8) is complex-valued for real μ , and the sign problem should arise when N is large.

We follow the steps given in Sect. 2.2 with the complexification of the variables $\phi_n \in (-\pi, \pi]$. We first set the largest flow time to $T = 2$, and prepare flow times $\{t_\alpha\}$ ($\alpha = 0, 1, 2, \dots, 20$) with equal separations as $t_0 = 0, t_1 = 0.1, t_2 = 0.2, \dots, t_{20} = 2.0$.⁸ Figure 1 shows $\phi(t) \equiv (1/N) \sum_n \phi_n(t)$ at flow times $t = t_\alpha$ with the $\phi_n(0)$ set to the same value. As Step 1 of our algorithm, we make a cold start ($\phi_n = 0$) for every replica α , and numerically solve the differential equations (2.7) and (2.8) with the adaptive 4th-order Runge–Kutta method to obtain the triplet $(x_\alpha, z_{t_\alpha}, J_{t_\alpha})$. We repeat the Metropolis process twenty times in Step 2,⁹ which is followed by a single sequence of swapping of Step 3. We then repeat Steps 2 and 3 ten times (as Step 4). With the first 5 data points discarded as initial sweeps, we estimate correlation functions with 10^4 data points. Figure 2 shows the absolute value of the denominator in Eq. (2.14) (divided by the sample size) as a function of μ with the other parameters set to $N = 8, m = 1, g^2 = 1/2$. The blue points are the result for the flow time $T = 0$. They correspond to the usual reweighting calculus, and show that the sign problem actually exists for $\mu \gtrsim 1.0$. The green (red) points are the result for the flow time $T = 2$ without (with) the parallel tempering implemented. The results show that the sign problem disappears at $T = 2$.

Figure 3 shows the chiral condensate $\langle \bar{\chi} \chi \rangle$ as a function of μ . The other parameters are again set to $N = 8, m = 1, g^2 = 1/2$. The dotted line represents the analytic result (3.7). The blue points are the result for the flow time $T = 0$ and exhibit large statistical errors, reflecting the sign problem. The green points are the result for the flow time $T = 2$ without the parallel tempering implemented.

⁸ The fermion determinant at flow time $t = 0$ is given by $\det D = (1/2^N) [\zeta + \zeta^{-1} + (\sqrt{2} + 1)^N + (\sqrt{2} - 1)^N]$ where $\zeta \equiv e^{i \sum_n \phi_n + N\mu}$. Note that $\det D$ vanishes at $\sum_n \phi_n = \pi \pmod{2\pi}$ when $\mu = \log(\sqrt{2} + 1) \sim 0.881$, which gives a multimodal distribution even at $t = 0$. However, this can be handled without introducing another tempering parameter, because each mode has a rather wide distribution (with not-too-small values near the boundary) and thus the whole configuration space can be easily explored by setting the interval of proposal distribution to be a large value, as in this paper.

⁹ As a proposal distribution we use the uniform distribution within the interval $[-\epsilon, \epsilon]$, where ϵ is chosen randomly from $\{1, 10^{-1}, 10^{-2}, \dots, 10^{-[2t_\alpha+1]}\}$ ($[k]$ is the floor of k).

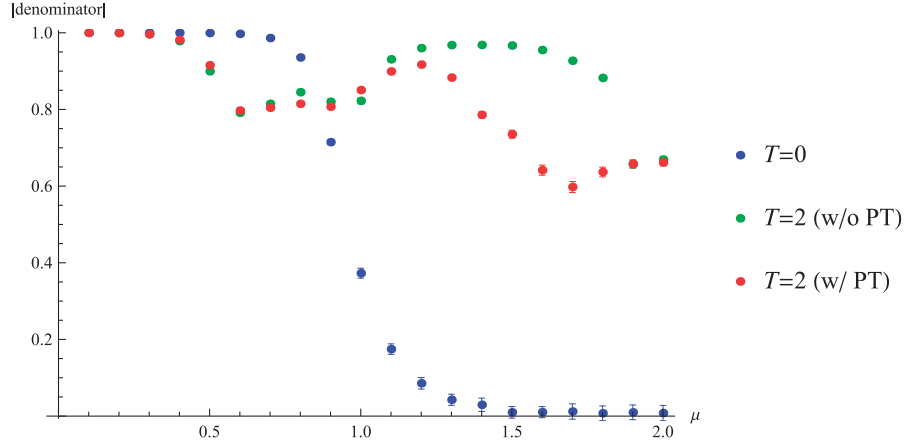


Fig. 2. The absolute value of the denominator in Eq. (2.14) (divided by the sample size) as a function of μ with the other parameters set to $N = 8$, $m = 1$, $g^2 = 1/2$. The blue points are the result for the flow time $T = 0$ and show that the sign problem actually exists for $\mu \gtrsim 1.0$. The green (red) points are the result for the flow time $T = 2$ without (with) the parallel tempering (PT) implemented, and show that the sign problem disappears at $T = 2$.

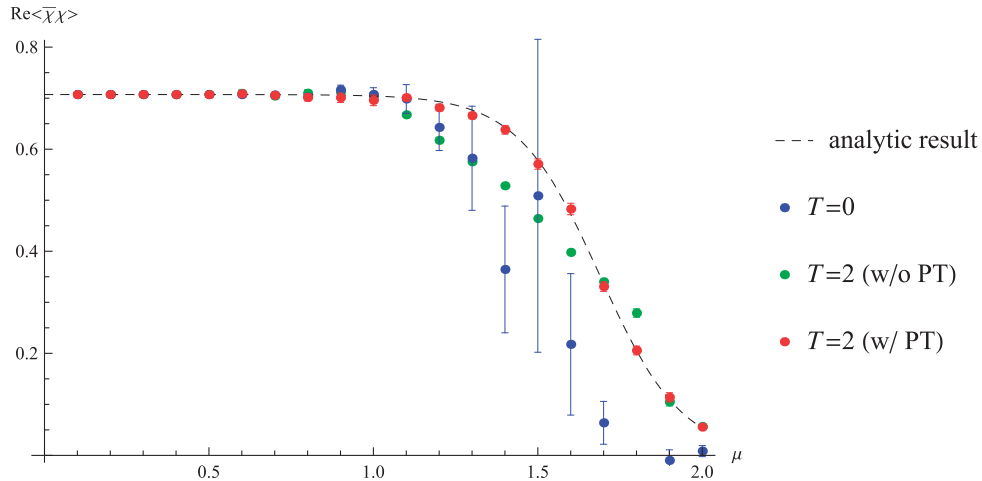


Fig. 3. Chiral condensate $\langle \bar{\chi}\chi \rangle$ as a function of μ with the other parameters set to $N = 8$, $m = 1$, $g^2 = 1/2$. The dotted line represents the analytic result (3.7). The blue (green) points are the result for the flow time $T = 0$ ($T = 2$) without the parallel tempering implemented. The red points are the result for the flow time $T = 2$ with the parallel tempering implemented.

They have small statistical errors, but exhibit statistically significant discrepancies from the analytic result. This should be attributed, as discussed in detail in Ref. [11], to the fact that the dominant contributions come only from a single thimble for such large T . The red points are the result for the flow time $T = 2$ now with the parallel tempering implemented. They show a good agreement with the analytic result, which implies that contributions from various thimbles are correctly taken into account through the parallel tempering. This can be confirmed by Figs. 4 and 5. Figure 4 exhibits the histogram of $\phi = (1/N) \sum_n \phi_n$ at $T = 2$. We see that the configurations are concentrated on a single thimble if the parallel tempering is not implemented (left), while they are spread out on various thimbles if the parallel tempering is implemented (right). Figure 5 shows the average acceptance

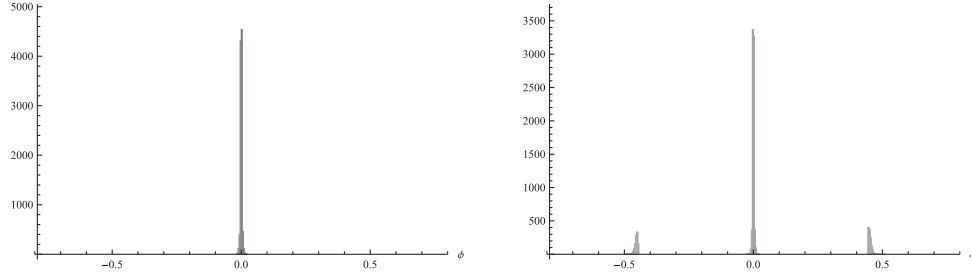


Fig. 4. Histograms of $\phi = (1/N) \sum_n \phi_n$ at the flow time $T = 2$ without/with the parallel tempering implemented (left/right). The parameters are set to $N = 8$, $m = 1$, $\mu = 1.3$, $g^2 = 1/2$.

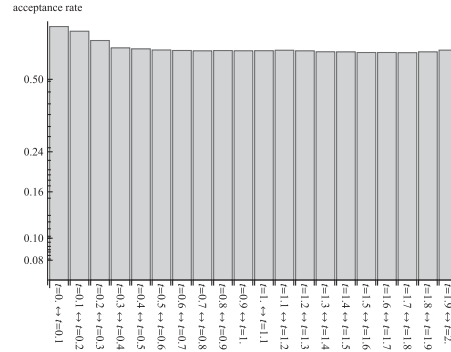


Fig. 5. Average acceptance rate for the swap between replicas α and $\alpha + 1$ in Step 3. The parameters are set to $N = 8$, $m = 1$, $\mu = 1.3$, $g^2 = 1/2$.

rates for the swaps between replicas α and $\alpha + 1$ in Step 3. We see that the swapping is carried out very well because the average acceptance rate is more than 50% for all pairs $(\alpha, \alpha + 1)$.

4. Conclusion and outlook

In this paper, as a versatile tool for Monte Carlo calculations of models with complex actions, we have proposed a Lefschetz thimble algorithm where the flow time is used as an auxiliary variable for the highly multimodal distribution. In particular, we implemented the parallel tempering method by taking the flow time t as a tempering parameter. There, we prepare flow times $\{t_\alpha\}$ such that $t_0 = 0$ and $t_A = T$. The largest flow time T can be taken to be sufficiently large such that the sign problem disappears there. Although the algorithm includes configurations at $t_0 = 0$, the original sign problem at $t_0 = 0$ does not enter the calculation because the sample average is taken only with respect to $t_A = T$, where the sign problem disappears. We have investigated the $(0+1)$ -dimensional massive Thirring model at finite density to exemplify that the algorithm does work, and showed that contributions from multi thimbles are correctly taken into account even for such a large flow time as $T = 2$.

We should investigate to what extent this algorithm is actually versatile. One interesting class of models, for which we can readily test our algorithm before applying it to QCD at finite density, is that of various types of large- N random matrix models with complex actions. In fact, if the free energy is calculated by an integration over matrices themselves (not over their eigenvalues), the classical solutions and the corresponding Lefschetz thimbles do not have a useful meaning because the quantum corrections are of the same order as the leading term. It thus provides us with a good test of versatility to check whether correct results are obtained for such models where the thimble

structure or its usefulness is not clear. As a related model, the numerical study of the triangle–hinge model [19–21] should also be interesting. The model is a sort of matrix model that generates 3D random volumes as a collection of triangles and hinges. In order to restrict the resulting configurations to tetrahedral decompositions, one needs to introduce a special form of interaction [19], which makes the action complex-valued (M. Fukuma, S. Sugishita, and N. Umeda, manuscript in preparation). A numerical study was made for a simplified model (with no restriction to tetrahedral decompositions), and the existence of a third-order phase transition is confirmed (manuscript in preparation). It is thus interesting to see whether the phase transition still exists when the restriction is imposed.

Besides the Lefschetz thimble algorithm, the complex Langevin algorithm [22,23] is also under intense study as a promising method to solve the sign problem. Recently, a very interesting proposal was made by Bloch [24] (see also Ref. [25]) to evaluate correlation functions by reweighting complex Langevin trajectories using such parameters that satisfy known validity conditions [26,27] to be free from wrong convergence problems [26–30]. It should be interesting to compare the extent of versatility between the reweighted complex Langevin algorithm and our parallel tempering algorithm with the flow time as a tempering parameter.

A study along these lines is now in progress and will be reported elsewhere.

Acknowledgements

The authors thank Hikaru Kawai, Jun Nishimura, Sotaro Sugishita, and Asato Tsuchiya for useful discussions. This work was partially supported by MEXT (Grant No. 16K05321).

Funding

Open Access funding: SCOAP³.

References

- [1] G. Aarts, J. Phys.: Conf. Ser. **706**, 022004 (2016) [arXiv:1512.05145 [hep-lat]] [Search INSPIRE].
- [2] M. Cristoforetti, F. D. Renzo, and L. Scorzato [AuroraScience Collaboration], Phys. Rev. D **86**, 074506 (2012) [arXiv:1205.3996 [hep-lat]] [Search INSPIRE].
- [3] M. Cristoforetti, F. Di Renzo, A. Mukherjee, and L. Scorzato, Phys. Rev. D **88R**, 051501 (2013) [arXiv:1303.7204 [hep-lat]] [Search INSPIRE].
- [4] A. Mukherjee, M. Cristoforetti, and L. Scorzato, Phys. Rev. D **88**, 051502(R) (2013) [arXiv:1308.0233 [physics.comp-ph]] [Search INSPIRE].
- [5] H. Fujii, D. Honda, M. Kato, Y. Kikukawa, S. Komatsu, and T. Sano, J. High Energy Phys. **1310**, 147 (2013) [arXiv:1309.4371 [hep-lat]] [Search INSPIRE].
- [6] M. Cristoforetti, F. Di Renzo, G. Erucci, A. Mukherjee, C. Schmidt, L. Scorzato, and C. Torrero, Phys. Rev. D **89**, 114505 (2014) [arXiv:1403.5637 [hep-lat]] [Search INSPIRE].
- [7] A. Alexandru, G. Başar, and P. Bedaque, Phys. Rev. D **93**, 014504 (2016) [arXiv:1510.03258 [hep-lat]] [Search INSPIRE].
- [8] T. Kanazawa and Y. Tanizaki, J. High Energy Phys. **1503**, 044 (2015) [arXiv:1412.2802 [hep-th]] [Search INSPIRE].
- [9] Y. Tanizaki, Y. Hidaka, and T. Hayata, New J. Phys. **18**, 033002 (2016) [arXiv:1509.07146 [hep-th]] [Search INSPIRE].
- [10] H. Fujii, S. Kamata, and Y. Kikukawa, J. High Energy Phys. **1511**, 078 (2015); **1602**, 036 (2016) [erratum] [arXiv:1509.08176 [hep-lat]] [Search INSPIRE].
- [11] A. Alexandru, G. Başar, P. F. Bedaque, G. W. Ridgway, and N. C. Warrington, J. High Energy Phys. **1605**, 053 (2016) [arXiv:1512.08764 [hep-lat]] [Search INSPIRE].
- [12] A. Alexandru, G. Başar, P. F. Bedaque, G. W. Ridgway, and N. C. Warrington, Phys. Rev. D **95**, 014502 (2017) [arXiv:1609.01730 [hep-lat]] [Search INSPIRE].
- [13] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **57** 2607 (1986).

- [14] C. J. Geyer, Markov chain Monte Carlo maximum likelihood, in *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*, ed. E.M. Keramidas (Interface Foundation, Fairfax Station, VA, 1991), p. 156.
- [15] D. J. Earl and M. W. Deem, Phys. Chem. Chem. Phys. **7**, 3910 (2005).
- [16] J. M. Pawłowski and C. Zielinski, Phys. Rev. D **87**, 094503 (2013) [arXiv:1302.1622 [hep-lat]] [Search INSPIRE].
- [17] J. M. Pawłowski, I. O. Stamatescu, and C. Zielinski, Phys. Rev. D **92**, 014508 (2015) [arXiv:1402.6042 [hep-lat]] [Search INSPIRE].
- [18] P. Hasenfratz and F. Karsch, Phys. Lett. B **125**, 308 (1983).
- [19] M. Fukuma, S. Sugishita, and N. Umeda, J. High Energy Phys. **1507**, 088 (2015) [arXiv:1503.08812 [hep-th]] [Search INSPIRE].
- [20] M. Fukuma, S. Sugishita, and N. Umeda, Prog. Theor. Exp. Phys. **2016**, 053B04 (2016) [arXiv:1504.03532 [hep-th]] [Search INSPIRE].
- [21] M. Fukuma, S. Sugishita, and N. Umeda, Prog. Theor. Exp. Phys. **2016**, 073B01 (2016) [arXiv:1603.05199 [hep-th]] [Search INSPIRE].
- [22] G. Parisi, Phys. Lett. B **131**, 393 (1983).
- [23] G. Aarts, L. Bongiovanni, E. Seiler, D. Sexty, and I. O. Stamatescu, Eur. Phys. J. A **49**, 89 (2013) [arXiv:1303.6425 [hep-lat]] [Search INSPIRE].
- [24] J. Bloch, Phys. Rev. D **95**, 054509 (2017) [arXiv:1701.00986 [hep-lat]] [Search INSPIRE].
- [25] J. Bloch, J. Glesaaen, O. Philipsen, J. Verbaarschot, and S. Zafeiropoulos, EPJ Web Conf **137**, 07030 (2017) [arXiv:1612.04621 [hep-lat]] [Search INSPIRE].
- [26] G. Aarts, F. A. James, E. Seiler, and I. O. Stamatescu, Eur. Phys. J. C **71**, 1756 (2011) [arXiv:1101.3270 [hep-lat]] [Search INSPIRE].
- [27] K. Nagata, J. Nishimura, and S. Shimasaki, Phys. Rev. D **94**, 114515 (2016) [arXiv:1606.07627 [hep-lat]] [Search INSPIRE].
- [28] T. Hayata, Y. Hidaka, and Y. Tanizaki, Nucl. Phys. B **911**, 94 (2016) [arXiv:1511.02437 [hep-lat]] [Search INSPIRE].
- [29] Y. Abe and K. Fukushima, Phys. Rev. D **94**, 094506 (2016) [arXiv:1607.05436 [hep-lat]] [Search INSPIRE].
- [30] L. L. Salcedo, Phys. Rev. D **94**, 114505 (2016) [arXiv:1611.06390 [hep-lat]] [Search INSPIRE].